

An iterative inversion of back-scattered acoustic waves

Ronan LeBras* and Robert W. Clayton*

ABSTRACT

The application of the Born approximation to the scattered wave field, followed by a WKBJ and far-field approximation on the propagation Green's function for a slowly varying background medium, leads to a simple integral relation between the density and bulk-modulus anomalies superimposed on the background medium and the scattered wave field. An iterative inversion scheme based on successive back-projections of the wave field is used to reconstruct the two acoustic parameters. The scheme, when applied to data generated using the direct integral relation, shows that the variations of the parameters can be reconstructed. The procedure is readily applicable to actual data, since every iterative step is essentially a prestack Kirchhoff migration followed by the application of the direct Born approximation and far-field operator.

INTRODUCTION

In recent years, several methods for inferring the nature of reflectors through the impedance contrasts causing the reflections have been presented in the literature. Those methods are referred to as inversions as opposed to the more commonly used migrations, since they aim at estimating the parameters of the medium rather than at describing the image of a reflector.

Bleistein and Cohen (1982) considered the acoustic wave equation for a uniform medium and introduced a velocity perturbation into the differential equation. They deduced an integral relation between the perturbation and the scattered wave field and they applied their method to stacked common-midpoint (CMP) data, losing the offset information in the process. Raz (1981) pointed out that to separate bulk modulus from density variations, it is necessary to have data from several offsets. Knowing both the density and the bulk modulus could prove to be a powerful way of better identifying the

materials composing the underlying medium (Gardner et al., 1974). Ostrander (1984) showed that the variation with offset of the amplitude of a reflected wave yields valuable information about the Poisson's ratio of material properties giving rise to anomalously high-amplitude reflections (bright spots). Ostrander applied his ideas to data recorded in an oil field and was able to discriminate convincingly between a gas-saturated rock layer and a basalt layer by comparing their Poisson's ratios.

Clayton and Stolt (1981) proposed a method that fully exploits the offset information contained in standard data. It is based on a Born model for the reflected field. Clayton and Stolt used the exact equation for an earth section that varies in two dimensions and introduced perturbations in both density and bulk modulus. Their approach is an attempt to retrieve more information and separate the two variables density and bulk modulus. For a multioffset survey, it becomes theoretically possible to invert the data for the relative variations of these two variables with depth and distance along the recording line or midpoint. Their approach to the inversion problem operates in the frequency-wavenumber domain. The wave field under the Born approximation can be expressed as a linear combination of the acoustic parameters of density and bulk modulus, a combination obtained in the midpoint-offset and depth-wavenumber spaces.

Under the Born approximation, the relationship between the parameters and the data becomes linear. (In the full theory the relationship is not linear.) Linear inverse theory can, for the Born approximation, be applied to the problem. However, the large amount of seismic data generally precludes the application of a generalized inverse theory to the problem. A typical problem in seismic reflection application is the imaging of a region having dimensions of a few tens of kilometers in length and several kilometers in depth. In order to apply the inverse theory, the medium has to be discretized. The smallest period present in the seismograms will determine the discretization step in space, if we want to take full advantage of the data. Usually this requirement gives a discretization step of 20 to 50 m, resulting in a dimension for the spatial variables of at least 10^6 . The dimensions of the data space are even

Presented at the 54th Annual International Meeting, Society of Exploration Geophysicists, 1984. Manuscript received by the Editor August 27, 1986; revised manuscript received September 4, 1987.

*California Institute of Technology, Seismological Laboratory 252-21, Pasadena, CA 91125.

©1988 Society of Exploration Geophysicists. All rights reserved.

larger. If we consider a survey, or a portion of a survey, consisting of 100 shots recorded by 24 receivers for a duration of 4 s and a discretization of 0.004 s, the dimension of the data space is 10^6 to 10^7 . Clearly, with such large dimensions for both model and data spaces, a classical least-squares method is impractical.

These considerations call for other techniques to search for the solution. A similar problem of large dimensions of data and model spaces was encountered by workers in geophysical tomography. The general problem in geotomography is estimating a velocity model from the knowledge of traveltimes residuals of seismic waves through the model, assuming a known travel path. (This last approximation can be likened to our assumption of knowing the background velocity.) Geophysical tomography has been applied to different data sets, including the whole earth, the regional crustal and Moho structures in Southern California (Hearn and Clayton, 1986a, b), the structure of the mantle under Southern California (Humphreys et al., 1984), and a local structure in France (Nercessian et al., 1984; Tarantola and Nercessian, 1984). Closer to our application, Tarantola (1984a, b) has introduced the concept of tomography for whole wave-field inversion of seismic data. All these authors use an iterative process based on back-projection to approximate the inverse. The back-projection operation consists of applying the transposed operator, or a modified version of that operator, to the residuals or to the data. The result is added to the previous model estimate to obtain an updated estimate. The next set of residuals is then calculated, and another iteration performed if desired. The iterative process allows problems with large dimensions to be handled by relatively small computers, since the data can be treated sequentially without requiring a large memory space corresponding to large matrices. Furthermore, no direct matrix inversion, with its inherent instabilities, is necessary.

One drawback to all the methods cited above is that the resolution and variance of the models obtained are difficult to estimate other than empirically, for example, by doing synthetic tests to estimate point-spread functions. In a following section, we introduce several optimization methods of a function, in this case the quadratic function ℓ_2 norm of the residuals, and show how to apply algorithms developed in the field of optimization (Fletcher, 1980) to our specific problem.

THE FORWARD MODEL FOR THE BACK-SCATTERED FIELD

An acoustic medium can be characterized by its bulk modulus κ and density ρ . We consider each of these fields to have a slowly varying background component (ρ_0 and κ_0) and a rapidly varying perturbation $\Delta\rho$ and $\Delta\kappa$. The wave equation relating the pressure field P due to a line source in a two-dimensional (2-D) acoustic medium is

$$LP \equiv \left(\frac{\omega^2}{\kappa} + \nabla \cdot \frac{1}{\rho} \nabla \right) P = \delta(x - x_0) \delta(z - z_0) S(\omega), \quad (1)$$

where (x_0, z_0) is the source location, and $S(\omega)$ is the Fourier transform of the source-time function. A forward solution for the Green's function G of equation (1) is given by the Lippman-Schwinger equation

$$G = G_0 + G_0 V G, \quad (2)$$

where L_0 and G_0 denote the wave operator and Green's functions respectively for the background parameters, and V ($V \equiv L - L_0$) is defined as the scattering potential.

This equation is exact but nonlinear in G . The first Born approximation linearizes the equation under the assumption that the scattering potential is small:

$$G = G_0 + G_0 V G_0. \quad (3)$$

The scattering potential V is a linear operator which can be related to the parameters of the medium by

$$V = \omega^2 \left(\frac{1}{\kappa} - \frac{1}{\kappa_0} \right) + \nabla \cdot \left(\frac{1}{\rho} - \frac{1}{\rho_0} \right) \nabla \equiv \omega^2 \frac{a_1}{\kappa_0} + \nabla \cdot \frac{a_2}{\rho_0} \nabla, \quad (4)$$

where we have followed the notation of Clayton and Stolt (1981) and introduced the perturbation parameters

$$a_1 = \frac{\kappa_0}{\kappa} - 1 \equiv \frac{\Delta\kappa}{\kappa} \quad (5)$$

and

$$a_2 = \frac{\rho_0}{\rho} - 1 \equiv \frac{\Delta\rho}{\rho}. \quad (6)$$

Equation (3) can now be used as the basis for a model of a back-scattered wave field recorded in a surface seismic reflection profile. To carry out the process, we remove the direct wave from equation (3) (the isolated G_0 term) and place the sources and receivers at the surface positions of $(x_s, 0)$ and $(x_g, 0)$, respectively. Consequently, the model for the recorded data D becomes

$$D(x_g, x_s; \omega) = G_0(x_g, 0 | x, z; \omega) V(x, z; \omega) \times G_0(x, z | x_s, 0; \omega) S(\omega). \quad (7)$$

This relation between the field of scattered data and the scattering potential is linear. We can take advantage of this fact to invert the data for the parameters defining the rapid variations of the acoustic medium.

Using a far-field approximation of equation (7) and transforming this equation to the space and time domains, we can put the total scattered wave field under the form of an integral over space:

$$D(x_s, x_g, t) = \int_{\Omega} d\Omega A(x_s, x) A(x, x_g) \times \left[a_1(x) + a_2(x) \cos \theta \right] \dot{S}(t - t_1 - t_2). \quad (8)$$

Equation (8) provides a basis for interpreting the assumptions in the Born model and for determining the model's limitations. The total scattered wave field has been cast in the form of a space integral over a product of three terms. The scattering potential term, containing the variables a_1 and a_2 , describes the variations of the scattered wave with the angle between incident and scattered wave on a single point scatterer composed of a_1 and a_2 . The terms A and \dot{S} are the amplitude and phase terms for the propagation between the source and the scatterer and the scatterer and the receiver. The dot over S indicates the time derivative. t_1 is the propagation time between the source and the scatterer and t_2 is the propagation time between the scatterer and the receiver. Thus, the whole

scattered field is computed as a linear superposition of waves scattered at each point of the medium considered as an independent single point scatterer.

The limitations of this approximation to the scattered wave field can be analyzed from expression (8). The propagation term is a ray-geometrical approximation, so that it is valid in the far field of both the source and the receiver. Interactions between different scatterers in the medium are neglected, which means (among other things) that multiples cannot be modeled using this approximation. Each point in the medium is considered as a weak heterogeneity; thus, propagation in media with high contrasts is not appropriately described by the Born approximation.

ℓ_2 NORM INVERSION

The inversion of equation (8) can be cast as an ℓ_2 norm optimization problem. The unknown parameters of the media are represented by the model-space vector $\mathbf{c} = (a_1, a_2)^T$. The forward problem described by equation (8) then becomes a linear system of equations,

$$\mathbf{d} = \mathbf{B}\mathbf{c}, \quad (9)$$

where \mathbf{B} represents the Born operator and \mathbf{d} is the data vector. The objective is to minimize the ℓ_2 norm of the residual vector defined as $\mathbf{d} - \mathbf{B}\mathbf{c}$. The ℓ_2 norm is a quadratic function of \mathbf{c} and can be optimized by techniques such as descent methods and quasi-Newton methods, among others (e.g., Rao, 1977; Fletcher, 1980).

The descent methods establish a line of search for a minimizing point along the direction of the gradient of a function, which in our case is the ℓ_2 norm of the residuals. The gradient is simply $-\mathbf{B}^T(\mathbf{d} - \mathbf{B}\mathbf{c})$. To calculate the gradient, it is necessary to compute the forward Born approximation at the point where we estimate the gradient and then to compute the transposed Born problem over the residual. The steepest descent algorithm is then

$$\mathbf{g}_k = \mathbf{B}^T(\mathbf{d} - \mathbf{B}\mathbf{c}_k),$$

$$\mathbf{s}_k = \alpha_k \mathbf{g}_k,$$

and

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \mathbf{s}_k,$$

where \mathbf{g}_k is the gradient in the model space at point \mathbf{c}_k , and α_k is a scalar computed so as to minimize the ℓ_2 norm of the residual. The starting model \mathbf{c}_0 is usually taken to be 0.

Descent methods have disadvantages. They may have a slow rate of convergence and may converge asymptotically. Quasi-Newton methods require only the knowledge of the first derivatives, which are easily computed in our case by back-projection of the residuals. The inverse of the Hessian is approximated by a positive definite matrix which is updated at each iteration. Approximating the inverse and updating it presents the advantage of no matrix inversion, since we try to estimate the inverse of the Hessian directly.

Quasi-Newton methods can be viewed as modifications of descent methods. Using the same notations as introduced in the last paragraph, iteration k of a quasi-Newton method can

be defined as follows:

$$\mathbf{g}_k = \mathbf{B}^T(\mathbf{d} - \mathbf{B}\mathbf{c}_k),$$

$$\mathbf{s}_k = \alpha_k \mathbf{H}_k \mathbf{g}_k,$$

and

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \mathbf{s}_k,$$

where \mathbf{g}_k is the gradient in the model space at point \mathbf{c}_k and α_k is a scalar computed so as to minimize the ℓ_2 norm of the residual. \mathbf{H}_k is a positive definite matrix. The Davidon-Fletcher-Powell (DFP) method (Fletcher, 1980) is a particular quasi-Newton method. In the case of the DFP method, the matrix \mathbf{H}_k is updated at each iteration following the relation

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} - \frac{(\mathbf{H}_k \mathbf{y}_k)(\mathbf{H}_k \mathbf{y}_k)^T}{\mathbf{y}_k^T \mathbf{H}_k \mathbf{y}_k},$$

where $\mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$ is the difference between the gradients at two consecutive iteration points.

The form of the equation updating the \mathbf{H}_k matrix indicates that only vectors of model-space dimension (two per iteration) are necessary to have a complete representation of \mathbf{H}_k .

The DFP method has important properties, especially when it is used to minimize quadratic functions as in our problem. For quadratic functions, the DFP method terminates in, at most, m iterations, where m is the size of the model space. Furthermore, \mathbf{H}_k approximates the inverse of the Hessian matrix at the point of convergence. In our case, the Hessian matrix is $\mathbf{B}^T \mathbf{B}$.

To compute the gradient of the ℓ_2 norm in the model space in our application, we need to define the transposed Born operator. This is done by using the following definition for the scalar product of two vectors \mathbf{x} and \mathbf{y} , which are functions of a variable describing a domain V which can be multidimensional:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \int_V dV \mathbf{x}^T \mathbf{y}.$$

The transposed operator \mathbf{B}^T is then fully defined by

$$\langle \mathbf{d}, \mathbf{B}\mathbf{c} \rangle = \langle \mathbf{B}^T \mathbf{d}, \mathbf{c} \rangle,$$

where \mathbf{c} belongs to the model space and \mathbf{d} to the data space.

Using the definition of matrix \mathbf{B} and of the scalar product, we have

$$\begin{aligned} & \int dx_s \int dx_g \int dt D(x_s, x_g, t) \int_{\Omega} d\Omega A(x_s, x) \\ & \times A(x, x_g)(a_1 + a_2 \cos \theta) \dot{S}(t - t_1 - t_2) = \langle \mathbf{d}, \mathbf{B}\mathbf{c} \rangle. \end{aligned}$$

The order of integration can be rearranged:

$$\begin{aligned} & \int_{\Omega} d\Omega \int dx_s \int dx_g \int dt A(x_s, x) A(x, x_g) \\ & \times (a_1 + a_2 \cos \theta) \dot{S}(t - t_1 - t_2) D(x_s, x_g, t) = \langle \mathbf{B}^T \mathbf{d}, \mathbf{c} \rangle. \end{aligned}$$

Hence, the transposed operator acting on the data set $D(x_s, x_g, t)$ is defined by

$$\begin{aligned} \hat{a}_1 &= \int dx_s \int dx_g \int dt A(x_s, s) A(x, x_g) \\ & \times \dot{S}(t - t_1 - t_2) D(x_s, x_g, t) \end{aligned} \quad (10)$$

and

$$\hat{a}_2 = \int dx_s \int dx_g \int dt A(x_s, s) A(x, x_g) \times \dot{S}(t - t_1 - t_2) \cos \theta D(x_s, x_g, t). \quad (11)$$

In practice, the procedure for obtaining the transposed data is to crosscorrelate the derivative of the source function and the seismograms with a time delay corresponding to the travel-time between the source, the current point of integration, and the receiver. The crosscorrelation is then weighted by the appropriate coefficient (amplitude factor for geometrical spreading and angular factor) and added to the array representing the parameter at the current point of the model space. This first iteration of the inversion procedure is very similar to a prestack Kirchhoff migration (Jain and Wren, 1980).

The preconditioning matrix \mathbf{H}_0 can be viewed as a modification to the prestack Kirchhoff migration. The initial \mathbf{H}_0 can be taken, for instance, as the identity matrix or as the diagonal matrix whose elements are the inverse of the diagonal elements of $\mathbf{B}^T \mathbf{B}$. We found that the choice of \mathbf{H}_0 is an important factor in the initial rate of convergence. The preconditioning consists in dividing the results of the back-projection by the quantities

$$\bar{a}_1 = \int dx_s \int dx_g \int dt A^2(x_s, x) A^2(x, x_g) \delta(t - t_1 - t_2)$$

and

$$\bar{a}_2 = \int dx_s \int dx_g \int dt A^2(x_s, x) A^2(x, x_g) \cos^2 \theta \delta(t - t_1 - t_2),$$

which are computed only once at the beginning of the inversion procedure. In our case, this choice of \mathbf{H}_0 means a larger initial computational effort to estimate that preconditioning matrix. The increased effort is counterbalanced by a larger reduction in the rms value at the first iteration.

APPLICATION TO SYNTHETIC DATA

The method given above was tested on synthetic data. In the first step, the data were generated using a direct Born formulation which allows us to concentrate on the performance of the inversion algorithm itself.

The models used for generating the synthetic data are very simple. They consist of an array of 30 by 20 cells with each cell 0.05 km square. The top of the imaged area (e.g., Figure 1) is at a depth of 0.2 km. The data were generated using a sampling interval of 0.004 s. The background velocity represents a linear increase with depth $v(z) = 2.5 + 0.343z$. Synthetic seismograms were computed at 16 midpoints and 24 offsets for each midpoint. The midpoints were placed at intervals of two cells (0.1 km) and the increment between successive offsets was four cell sizes (0.2 km). This choice gave us a maximum offset of 4.6 km. The first midpoint was directly above the top left corner of the area to be imaged. The source function used for the examples presented in Figures 1 through 10 was one cycle of a sine function.

Point scatterers

Synthetic data were generated for point scatterers simulated by placing a cell with an anomaly either in density or in bulk

modulus at the center of the area to be imaged. In both cases, the amplitude of the anomaly was a relative increase of 0.01 in the variable at the point scatterer. Figure 1 shows the results of the first and the third iterations for synthetic data generated from a point scatterer with variable density. The intensity scale at the bottom was chosen so that the maximum amplitudes on the sections were slightly saturated. The top panels on Figures 1 and 2 represent the sections after one iteration of the inversion algorithm. Figure 2 is the section from a bulk modulus point scatterer and also presents the direct results of the back-projections of the first and third iterations (upper and middle panels), along with smoothed versions (bottom panels).

These two examples illustrate some of the characteristics of this inversion. Artifacts are similar to those of Kirchhoff migration, as could be expected. This similarity is best illustrated by the figures showing the first iteration of the inversion for the two different models. There is a streaking of the anomalies in the horizontal direction, which can be thought of as the effect of superimposing the equal-traveltime arcs. The arcs do not interfere destructively due to the absence of arcs crossing vertically. This behavior is very much improved with successive iterations.

Another feature of this iterative method is the buildup of the amplitude of the anomalies with successive iterations.

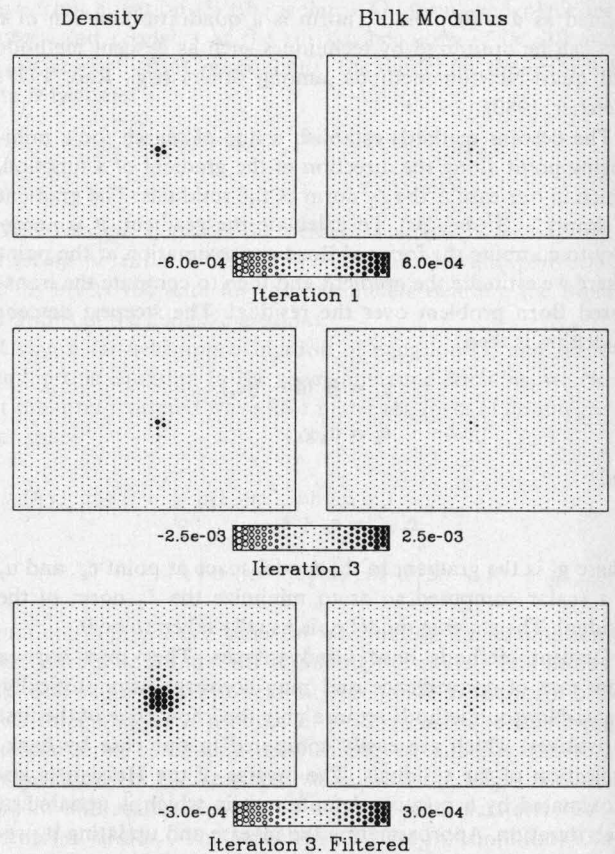


FIG. 1. Density point scatterer. The top two panels are the results of the first iteration; the middle panels are the results of the third iteration; and the lower panels, the filtered sections for the third iteration (smoothed over a disk of radius two cells).

Buildup is illustrated by the difference in the amplitude scales on the figures between the first and third iterations, reflecting the difference in the maximum amplitudes on the sections. This amplitude buildup is similar to that observed in tomographic inversion of traveltime data. Several problems were examined relative to this inversion method and are presented below.

To illustrate how such an inversion technique is useful in differentiating between different geometries, we computed synthetics for a model with four point scatterers. Each of the points has different characteristics. The model is presented in Figure 3. Three sets of reconstructions have been made for this model; first with a maximum offset of 1.15 km, and then with maximum offsets of 2.3 and 4.6 km. Figure 4 shows the reconstructions for the small maximum offset of 1.15 km; Figure 5 shows the reconstructions for a maximum offset of 2.3 km; and Figure 6, the reconstruction for a maximum offset of 4.6 km. The bulk modulus section was made up of three scatterers with equal weights of 0.01. The density section has anomalies of 0.01, 0.02, and -0.01 . Note that for the intermediate and large maximum offsets, the anomalies are fairly well reconstructed, except for the lower right corner anomaly, which is poorly recovered (although recovered better in the large-offset experiment, Figure 6). This poor reconstruction of the lower right anomaly is explainable, since that anomaly has no impedance contrast with the background and therefore will not radiate any backscattered energy in the direction inverse to the incident wave. The surface data, being mostly back-

scattered, will not see the anomaly, unless the offset coverage becomes such that the scattered waves are at least partially forwardly scattered. In all cases, the impedance anomalies are well reconstructed.

The main feature of the small-offset reconstruction (Figure 4) is the lack of discrimination between the different rheologies of the point scatterers. The density and bulk modulus sections are roughly similar, which means that the impedance is recovered but the offset is not sufficient to discriminate between different anomalies.

Noise analysis

An empirical noise analysis was conducted by adding white noise to a signal generated with a density point scatterer model. The results are very encouraging. The effect of adding white noise even in large amounts did not affect the back-projected images very much. This lack of sensitivity to noise is due to the ability of the method to cancel random noise by adding the noise destructively along equal-traveltime paths.

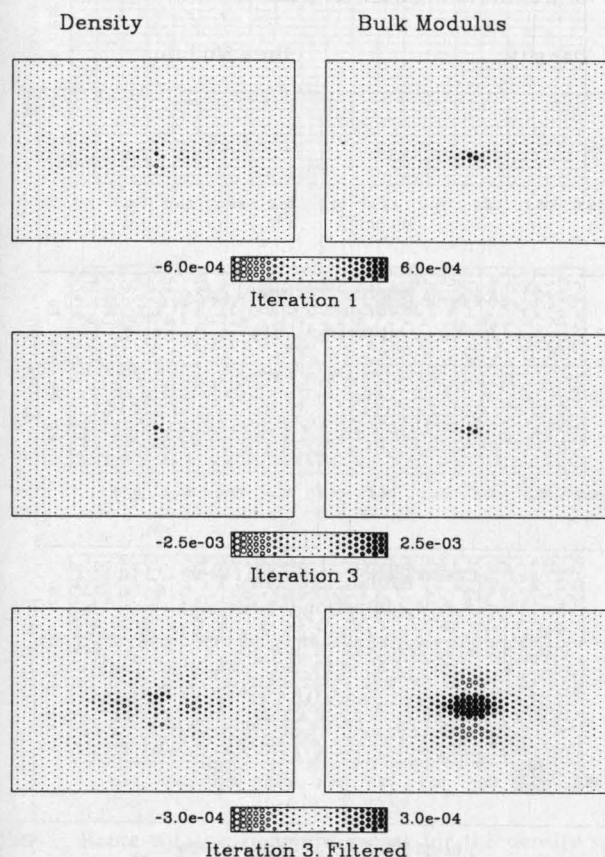


FIG. 2. Modulus point scatterer. The panels are in the same order as in Figure 1.

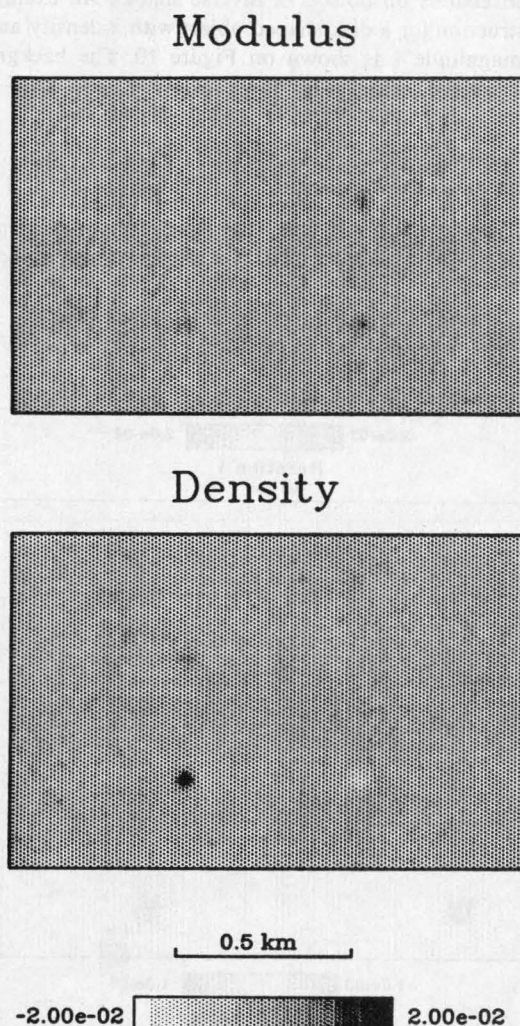


FIG. 3. Model consisting in four point scatterers. Each scatterer has different characteristics, as indicated by this figure (see text).

Figure 7 shows representative midpoint gathers for the one-point density scatterer, with signal-to-noise ratios of ∞ , 10, and 0.05. Figure 8 shows the direct results of the first iteration (top panel) and third iteration (middle panel) for a signal-to-noise ratio of 10 and also shows the smoothed section for the third iteration (bottom panel). These results can be compared with those of Figure 1 which show the noiseless reconstructions. Figure 9 shows the results of the first and third iterations, both unfiltered and filtered for a signal-to-noise ratio of 0.05. The effect of the noise is more readily apparent in this case, but even in extreme cases with very low signal-to-noise ratios, the method is robust in the presence of random noise. There is cancellation of noise along the equal-traveltime paths.

Disk-shaped anomaly

The examples presented so far have been point scatterers. Since the Born approximation leads to a linear relationship between the parameters of the medium and the synthetic seismograms, one can think of more complex models as linear superpositions of such point scatterers. Scattered waves then combine and interfere constructively to give rise to reflections and diffractions on objects of diverse shapes. An example of reconstruction for a disk-shaped object with a density anomaly of magnitude 1 is shown on Figure 10. The background medium is the same as for the previous examples of point

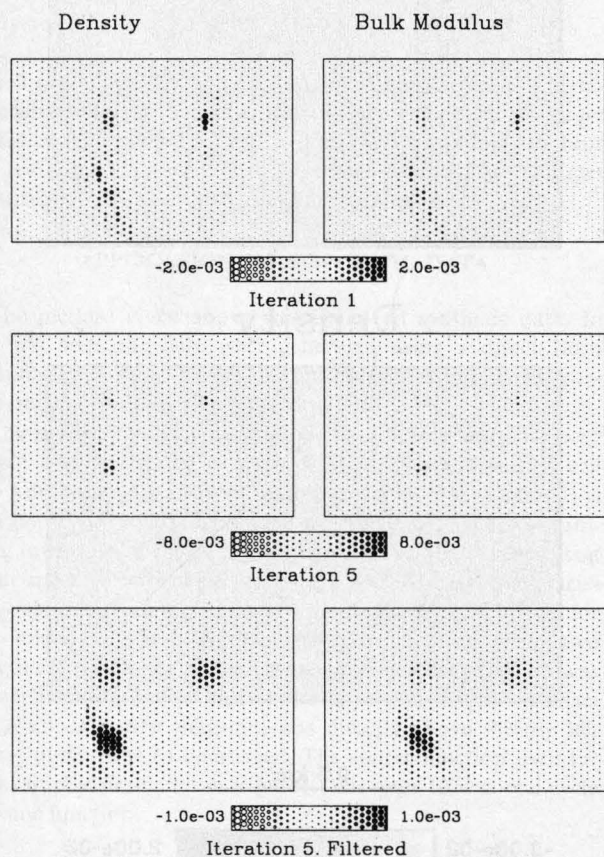


FIG. 4. Results of the inversion for the four point scatterers, with a maximum offset of 1.15 km. The top two panels are the results after one iteration; the middle two panels are the results for the fifth iteration; and the bottom two panels, the fifth iteration smoothed over a disk of radius two cells.

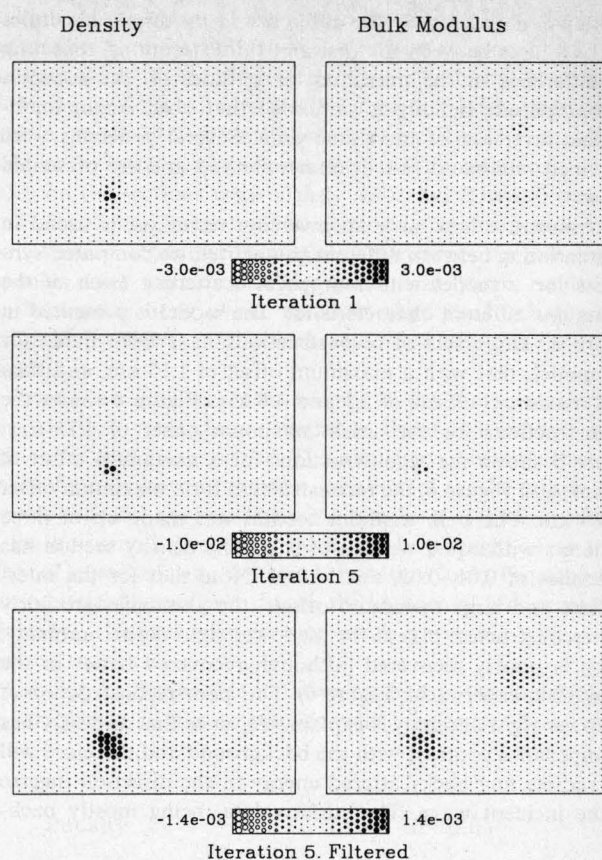


FIG. 5. Same as Figure 4, for a maximum offset of 2.3 km.

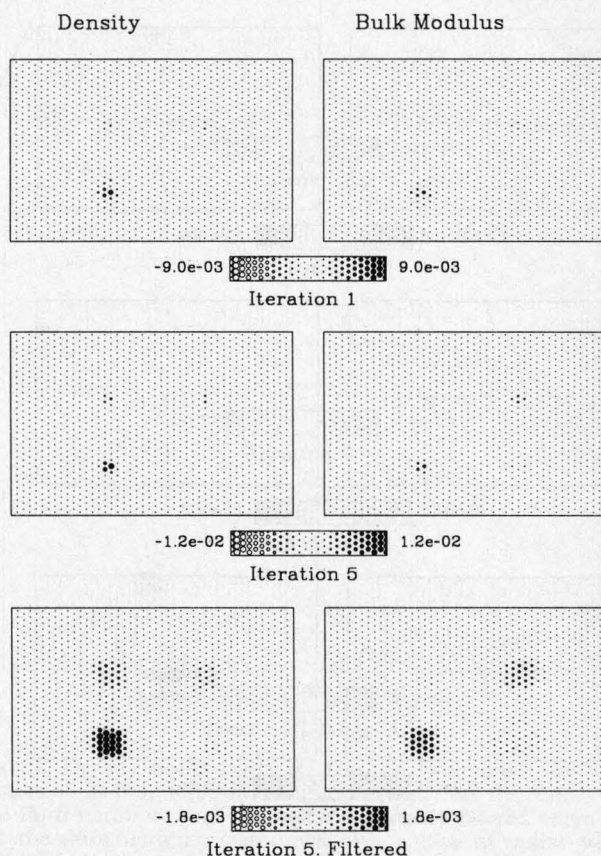


FIG. 6. Same as Figure 4, for a maximum offset of 4.6 km.

scatterers (linear increase with depth). The original disk has a diameter of 10 cells (0.5 km). Figure 10 displays iterations 1, 3, 5, and 7 of the inversion. The sections are all plotted at the same scale and the buildup in amplitude with iterations is more easily observed on this example. Note also the improvement in the streaking. The reductions in the rms value were, respectively, 58 percent, 69 percent, 80 percent, and 88 percent for these iterations.

CONCLUSIONS

A method has been presented for inverting seismic data to obtain the acoustic parameters of the subsurface. The method is based on a Born approximation for the acoustic wave field and an additional far-field approximation. It is applicable when the background acoustic parameters are slowly varying functions of space. The Green's functions can then be computed using a WKB approximation. The method is iterative, and each of its steps resembles a prestack Kirchhoff migration. The difference between this inversion and a prestack Kirchhoff migration is that coefficients multiply the back-projected field. These coefficients depend upon the geometrical spreading and the angle between the incident and the scattered waves.

The method was tested on synthetic data, and it was shown that a good image of the impedance can be obtained even after a few iterations. When the data contain large offsets, a differentiation between two independent acoustic parameters

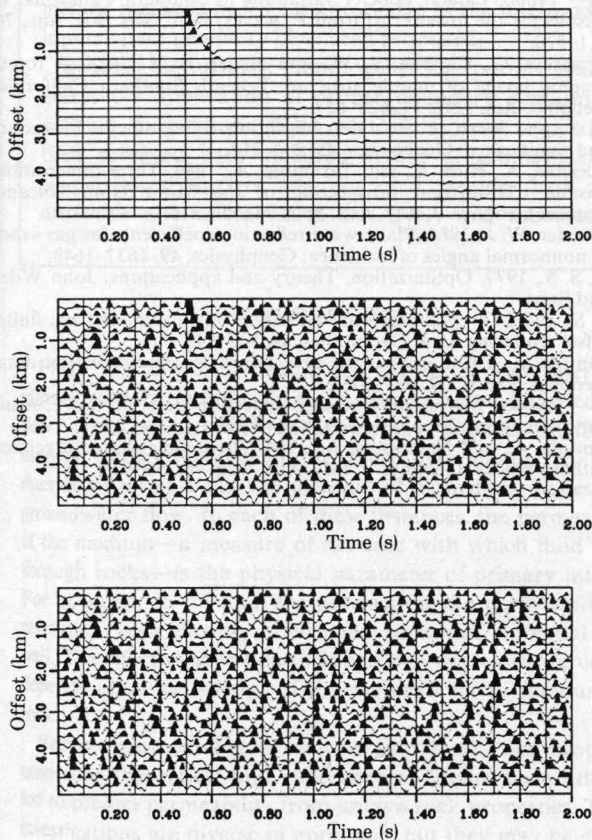


FIG. 7. Representative midpoint gather for the density scatterer model. The top section is for a signal-to-noise (S/N) ratio of ∞ ; middle section, S/N ratio of 10; and bottom section, for an S/N ratio of 0.05.

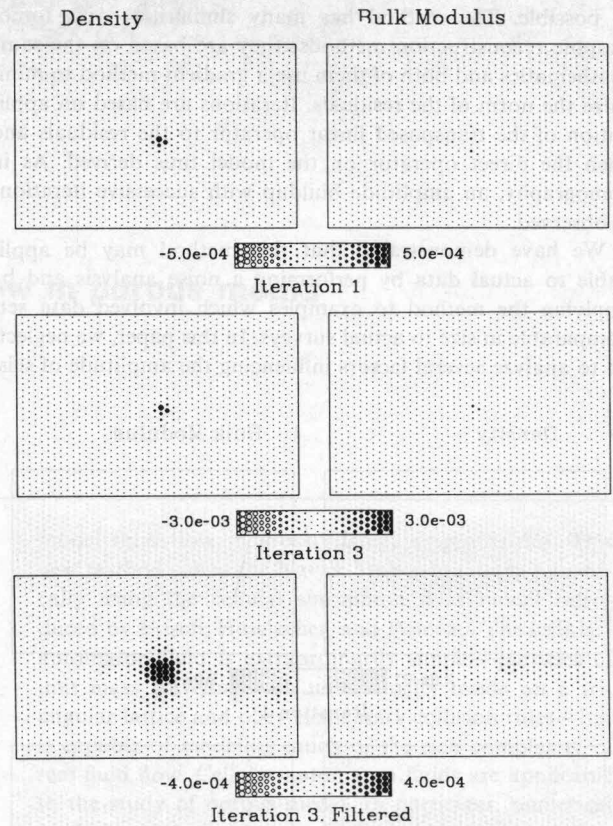


FIG. 8. Reconstructions for a signal-to-noise ratio of 10. The top two panels are the results of the first iteration; the middle panel, the third iteration; and the bottom panels, the smoothed sections of the third iteration.

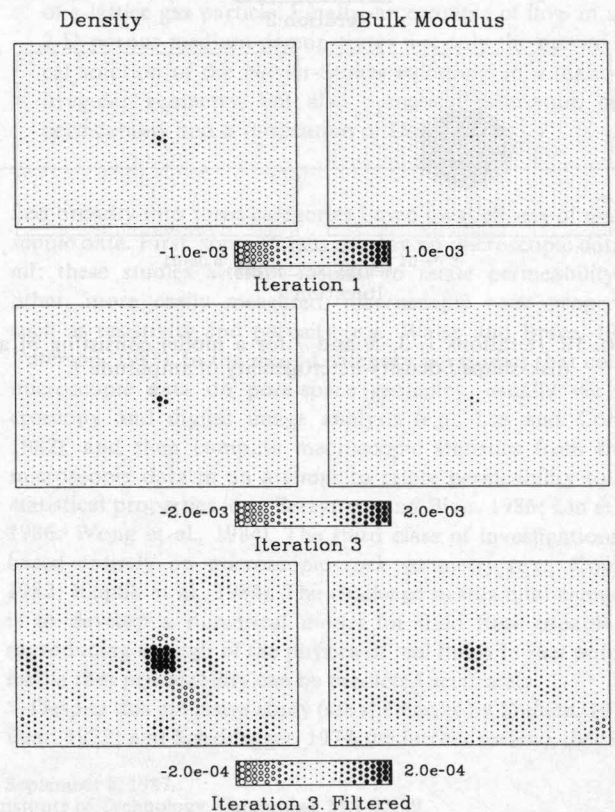


FIG. 9. Same as Figure 8 for a signal-to-noise ratio of 0.05.

is possible. The method has many similarities with tomographic reconstruction methods; they are based on the same mathematics and both of them use a gradient method to minimize the norm of the residuals. Iterations are based on application of the transposed linear operator to the residuals and then the direct operator on the model thus derived. As in tomography, an amplitude buildup with successive iterations is observed.

We have demonstrated that this method may be applicable to actual data by performing a noise analysis and by applying the method to examples which involved data sets comparable in size to actual surveys. In this paper, we neglected to analyze several factors influencing the amplitude of seismic data. These factors are, among others, the radiation pattern of the source, the effect of the free surface, the attenuation of the receiver array which is a function of the incidence angle on the array of receivers, and, perhaps most importantly, the anelastic attenuation of the medium. The problem with the radiation pattern from the source can be easily tackled with the high-frequency approximation approach used here.

ACKNOWLEDGMENTS

We would like to thank John Louie and Christof Stork from Caltech for their help and suggestions. We would also like to thank Prof. Albert Tarantola from I.P.G. (Paris) for suggesting the iterative method presented here. This work was partially supported by a National Science Foundation PYI award to RWC.

Contribution no. 4353 from the Division of Geological and Planetary Science, California Institute of Technology.

REFERENCES

- Bleistein, N., and Cohen, J. K., 1982, The velocity inversion problem. Present status, new directions: *Geophysics*, **47**, 1497-1511.
- Clayton, R. W., and Stolt, R., 1981, A Born-WKB inversion method for acoustic reflection data: *Geophysics*, **46**, 1559-1567.
- Fletcher, R., 1980, Practical methods of optimization: John Wiley and Sons.
- Gardner, G. H., Gardner, L. W., and Gregory, A. R., 1974, Formation velocity and density—the diagnosis basics for stratigraphic traps: *Geophysics*, **39**, 770-780.
- Hearn, T. M., and Clayton, R. W., 1986a, Lateral velocity variations in Southern California. I. Results for the upper crust from P_g waves: *Bull., Seis. Soc. Am.*, **76**, 495-509.
- 1986b, Lateral velocity variations in Southern California. II. Results for the lower crust from P_n waves: *Bull., Seis. Soc. Am.*, **76**, 511-520.
- Humphreys, G., Clayton, R. W., and Hager, B. H., 1984, A tomographic image of mantle structure beneath Southern California: *Geophys. Res. Lett.*, **11**, 625-627.
- Jain, S., and Wren, A. E., 1980, Migration before stack—Procedure and significance: *Geophysics*, **45**, 204-212.
- Nercissian, A., Hirn, A., and Tarantola, A., 1984, Three-dimensional seismic transmission prospecting of the Mont Dore volcano, France: *Geophys. J. Roy. Astr. Soc.*, **76**, 307-315.
- Ostrander, W. J., 1984, Plane-wave reflection coefficients for gas sands at nonnormal angles of incidence: *Geophysics*, **49**, 1637-1648.
- Rao, S. S., 1977, Optimization. Theory and applications: John Wiley and Sons.
- Raz, S., 1981, Three-dimensional velocity profile inversion from finite-offset scattering data: *Geophysics*, **46**, 837-842.
- Tarantola, A., 1984a, Linearized inversion of seismic reflection data: *Geophys. Prosp.*, **32**, 998-1015.
- 1984b, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259-1266.
- Tarantola, A., and Nercissian, A., 1984, Three-dimensional inversion without blocks: *Geophys. J. Roy. Astr. Soc.*, **76**, 299-306.

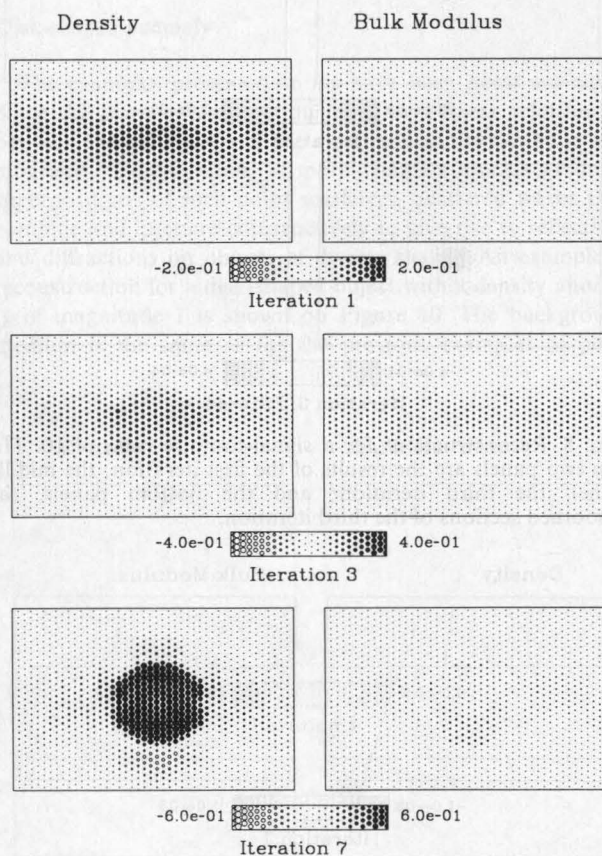


FIG. 10. Iterations 1, 3, 5, and 7 for a model consisting of a disk-shaped density heterogeneity of amplitude 1.